
Chapter 1: Introduction

1.1 The National Longitudinal Survey of Youth 1979 (NLSY79)

The NLSY79 is a nationally representative sample of 12,686 young men and women who were 14-22 years old when they were first surveyed in 1979. These individuals are now in their late thirties to mid forties, and have been personally interviewed for more than two decades. Since their first interview, many of the respondents have made the transitions from school to work, and from their parents' homes to becoming parents and homeowners. Data collected yearly from 1979 to 1994, and biennially from 1996 to the present, chronicle these changes and provide researchers an opportunity to study in great detail the experiences of a large group of young adults who can be considered representative of all American men and women born in the late 1950s and early 1960s.

The primary purpose of the NLSY79 is the collection of data on each respondent's labor force experiences, labor market attachment, and investments in education and training. However, the actual content of the NLSY79 is much broader due to the interests of governmental agencies besides the Department of Labor. At several points throughout the survey, various agencies have funded special sets of questions. In addition to major funding, supportive funding has been provided by the U.S. Department of Health and Human Services, the National Institute on Aging, the National Institute on Alcohol Abuse and Alcoholism, the U.S. Department of Defense and the Armed Services and the National Institute of Education. Examples of other topical areas include:

- (1) Military participation—Support from the Department of Defense and the Armed Services made possible the 1979–84 interviews of 1,280 youth enlisted in the military.
- (2) Vocational aptitude—The *Armed Services Vocational Aptitude Battery (ASVAB)*, a study which was jointly sponsored by the Departments of Defense and Labor, was administered to the civilian and military youth samples in 1980.
- (3) High school performance—Beginning in 1979, a five-year cooperative effort of the National Center for Research in Vocational Education and The Ohio State University's Center for Human Resource Research resulted in a survey of the high schools of civilian NLSY79 respondents and the collection of detailed transcript information on potential high school graduates.
- (4) Time-use—In 1981, the National Institute of Education sponsored a set of time-use questions.
- (5) Alcohol and substance use—Funding from the National Institute on Alcohol Abuse and Alcoholism and the National Institute on Drug Abuse has made possible multiple fieldings of expanded sets of alcohol and substance use questions.
- (6) Children's issues—Assessments of NLSY79 children, the development of the fertility and the childcare components of the youth surveys, and a child school survey were made possible through funding from the National Institute of Child Health and Human Development.
- (7) Department of Justice—A self-report supplement to the 1980 survey collected data on the respondent's participation in and income from various delinquent and criminal activities.

This detailed guide is designed for researchers who are either working or planning to work with the National Longitudinal Survey of Youth 1979 (NLSY79). Users who are interested in a general overview of the NLS surveys should refer to the *NLS Handbook*. Users may also wish to read recent issues of *NLS News*, a quarterly newsletter, available on-line at <http://www.bls.gov/nls/nlsnews.htm>. For those not familiar with the NLS, a brief overview of the NLSY79 and the other NLS cohorts follows.

1.2 The NLS

The National Longitudinal Surveys (NLS) are a set of surveys sponsored by the Bureau of Labor Statistics (BLS), U.S. Department of Labor. These surveys have gathered information at multiple points in time on the labor market experiences of diverse groups of men and women. Each of the NLS samples consists of several thousand individuals, some of whom have been surveyed over several decades. The earliest NLS interviews began in 1966 under the original sponsorship of the Office of Manpower, Automation, and Training (now the Employment and Training Administration). These cohorts were chosen in an effort to understand specific issues pertaining to the U.S. labor market, such as retirement, the return of housewives to the labor force, and the school-to-work transition. Since that time, however, the content of the surveys has been expanded to provide useful information on an extremely broad range of topics.

The first four NLS cohorts (Older Men, Mature Women, Young Men, and Young Women) were selected in the mid-1960s because each faced important labor market decisions that were of special concern to policy makers. Although the initial plan called for only a 5-year period of interviewing, high retention rates and widespread research interest led investigators to continue the surveys. In 1977, a survey of all known NLS data users and the recommendations of a panel of experts convened by the Department of Labor resulted in two decisions. The first was to continue the surveys of the four Original Cohorts for an additional five years (as long as attrition did not become a problem). The second decision was to begin a new longitudinal study of a panel of young men and young women.

This new study was initiated to permit a replication of the analysis of the 1960s Young Men and Young Women cohorts and to assist in the evaluation of the expanded employment and training programs for youth legislated by the 1977 amendments to the Comprehensive Employment and Training Act (CETA). To these ends, in 1978 a national probability sample was drawn of young women and young men living in the United States and born between January 1, 1957 and December 31, 1964. This sample included an overrepresentation of blacks, Hispanics, and economically disadvantaged non-black/non-Hispanics. With funding from the Department of Defense and the Armed Services, an additional group of young persons serving in the military was selected for interviewing. This sample of civilian and military youth, called the National Longitudinal Survey of Youth 1979 (NLSY79), was first interviewed in early 1979 and has been re-interviewed regularly since that date.

The Children of the NLSY79 survey, begun in 1986, further enhanced the NLS. With funding from the National Institute of Child Health and Human Development (NICHD) and a number of additional government agencies and private foundations, detailed information on the development of children born to NLSY79 women has supplemented the data on mothers and children collected during the regular youth surveys. During these biennial surveys, a battery of child cognitive, socio-emotional, and physiological assessments are administered to NLSY79 mothers and their children. In addition to these assessments, the Children of the NLSY79 are also asked a number of questions in an interview setting. In 1994, children age 15 and older, the “Young Adults,” first responded to a separate survey with questions similar to those asked of their mothers and a wide array of attitudinal and behavioral questions tailored to their age group.

With the aging of the NLSY79 cohort, another longitudinal cohort has been started. The National Longitudinal Survey of Youth 1997 (NLSY97) collects information on the circumstances that influence or are influenced by the labor market behaviors of youth ages 12–16 as of December 31, 1996. Data on the youths’ educational experiences, along with their family and community backgrounds, are also included in the survey. Documenting the transition from school to work, this survey is designed to be representative of the population born during the period 1980 to 1984.

Information on sample sizes, interview years, and the current survey status of each respondent group is presented in Table 1.2.1 below.

Table 1.2.1 The NLS: Survey Groups, Sample Sizes, Interview Years & Status

Survey Group	Age Cohort	Birth Year Cohort	Initial Sample Size	Initial/Latest Survey Year	# of Surveys to Date	Survey Status
Older Men	45–59 (as of 3/31/66)	4/1/06–3/31/21	5020	1966/1990	13 ¹	Ended
Mature Women	30–44 (as of 3/31/67)	4/1/22–3/31/37	5083	1967/2003	21	Continuing
Young Men	14–24 (as of 3/31/66)	4/1/41–3/31/52	5225	1966/1981	12	Ended
Young Women	14–24 (as of 12/31/67)	1/1/43–12/31/53	5159	1968/2003	22	Continuing
NLSY79	14–21 (as of 12/31/78)	1957–1964	12686	1979/2002	20	Continuing
NLSY79 Children	Birth–14	—	³	1986/2002	9	Continuing
NLSY79 Young Adults ²	15 & older	—	³	1994/2002	5	Continuing
NLSY97	12–16 (as of 12/31/96)	1980–1984	8984	1997/2003	6	Continuing

¹ Twelve interviews occurred from 1966–83. The 1990 interview surveyed living respondents and next-of-kin of deceased respondents.

² NLSY79 Young Adult respondents were initially interviewed as part of the NLSY79 Children sample. Beginning in 1994, those 15 and older were surveyed separately. In 1998, youths 21 and older were not interviewed and in 2000 the full age range was interviewed.

³ The sizes of the NLSY79 Children and Young Adult samples are dependent on the number of children born to NLSY79 respondents. Since this number is still increasing, original sample sizes are omitted.

Administration of the NLSY79 Project. Responsibility for the administration of the NLS resides with the Bureau of Labor Statistics (BLS), an agency of the U.S. Department of Labor. Established in 1884, BLS is responsible for the analysis and publication of data series on employment and unemployment, prices and living conditions, compensation and working conditions, productivity, occupational safety and health, and economic growth and employment projections. Its mission is to promote the development of the U.S. labor force by gathering information about the labor force and disseminating it to policy makers and the public so participants in those markets can make more informed, and thus more efficient, choices.

The NLS supports BLS in this mission. The surveys are part of a longitudinal research program that includes in-house analyses, extramural grants, and other special projects. The NLS program is housed in the Office of Employment Research and Program Development, headed by an Assistant Commissioner. Direction of the NLS program is the responsibility of the Director of National Longitudinal Surveys.

BLS contracts with the Center for Human Resource Research (CHRR) at The Ohio State University to manage the NLS Original Cohorts and NLSY79, to share in the design of the survey instruments, to disseminate the data, and to provide user services. The Census Bureau collects data for the Original Cohorts through an interagency agreement with BLS. Data collection for the NLSY79 and the NLSY79 Children samples is subcontracted to the National Opinion Research Center (NORC) at the University of Chicago. BLS contracts with NORC to manage the NLSY97 cohort and to collect the data for this survey; CHRR is subcontracted for instrument programming, data dissemination, documentation, and variable creation. The project is assisted in its efforts by the NLS Technical Review Committee. Meeting twice each year, committee members provide recommendations regarding questionnaire design, additional survey topics, potential research uses, methodological issues, data distribution, and user services. The committee is multidisciplinary, reflecting the wide range of social scientists using NLS data.

Ultimate responsibility rests with BLS for overseeing all aspects of the work undertaken by the other organizations. For more information about the NLS program, contact:

The National Longitudinal Surveys
Bureau of Labor Statistics
2 Massachusetts Avenue NE, Room 4945
Washington, D.C. 20212-0001
(202) 691-7410; fax: (202) 691-7425
e-mail: NLS_info@bls.gov
www: <http://www.bls.gov/nls>

1.3 NLSY79 Samples

The NLSY79 sampling design enables researchers to analyze the experiences of groups such as women, Hispanics, blacks, and the economically disadvantaged. The following three subsamples comprise the NLSY79:

- (1) a cross-sectional sample of 6,111 respondents designed to be representative of the non-institutionalized civilian segment of young people living in the United States in 1979 and born between January 1, 1957, and December 31, 1964 (ages 14–21 as of December 31, 1978)
- (2) a supplemental sample of 5,295 respondents designed to oversample civilian Hispanic, black, and economically disadvantaged non-black/non-Hispanic youth living in the United States during 1979 and born between January 1, 1957, and December 31, 1964
- (3) a sample of 1,280 respondents designed to represent the population born between January 1, 1957, and December 31, 1961 (ages 17–21 as of December 31, 1978), and who were enlisted in one of the four branches of the military as of September 30, 1978

Users can identify a respondent's sample type by using variable R01736. With a few exceptions, all members of the cross-sectional sample have been eligible for interview during each NLSY79 survey. Funding constraints imposed limitations on the numbers of military and supplemental sample members who were eligible for interview after the 1984 and 1990 surveys, respectively. Following the 1984 interview, 1,079 members of the military subsample were no longer eligible for interview; 201 respondents randomly selected from the entire military subsample remained in the survey. Following the 1990 interview, none of the 1,643 members of the economically disadvantaged, non-black/non-Hispanic subsample were eligible for interview. Table 1.3.1 presents sample sizes for several survey years.

User Notes: Although the entire economically disadvantaged, non-black/non-Hispanic subsample was dropped following the 1990 survey, the 'Reason for Noninterview' variable has only 1,621 respondents listed as "supplemental poor white sample dropped." The remaining 22 respondents died prior to the dropping of the subsample. These individuals continue to be classified as "deceased." For further information, refer to the "Reasons for Noninterview" section in chapter 2.

Table 1.3.1 NLSY79 Sample Size by Subsample, Race & Gender for Selected Years

Year	1979	1984	1990	1994	1996	1998	2000	2002
Total Sample	12686	12069	10436	8891	8636	8399	8033	7724
Cross-Sectional Sample	6111	5814	5498	5457	5290	5159	4949	4775
Males	3003	2839	2664	2648	2554	2459	2356	2270
Non-black/non-Hispanic	2439	2303	2157	2150	2085	2007	1926	1852
Black	346	329	318	310	301	289	272	268
Hispanic	218	207	189	188	168	163	158	150
Females	3108	2975	2834	2809	2736	2700	2593	2505
Non-black/non-Hispanic	2477	2365	2271	2243	2177	2161	2065	1999
Black	405	393	365	363	358	346	343	328
Hispanic	226	217	198	203	201	193	185	178
Supplemental Sample	5295	5040	4755	3256	3171	3065	2921	2792
Males	2576	2442	2280	1599	1555	1480	1412	1353
Poor non-black, non- Hispanic	742	699	664 ¹	—	—	—	—	—
Black	1105	1055	979	973	946	883	847	828
Hispanic	729	688	637	626	609	597	565	524
Females	2719	2598	2475	1657	1616	1585	1509	1440
Poor non-black/non-Hispanic	901	851	819 ¹	—	—	—	—	—
Black	1067	1034	984	987	971	951	912	879
Hispanic	751	713	672	670	645	634	597	561
Military Sample	1280	1215	183²	178	175	175	163	157
Males	824	774	168	164	160	161	152	147
Non-black/non-Hispanic	609	575	70	68	68	66	64	61
Black	162	151	68	62	61	63	61	60
Hispanic	53	48	30	34	31	32	27	26
Females	456	441	15	14	15	14	11	10
Non-black/non-Hispanic	342	331	7	7	7	6	6	6
Black	89	86	5	5	5	5	3	3
Hispanic	25	24	3	2	3	3	2	1

¹ Subsample dropped after the 1990 interview.

² Two hundred and one members of the military sample were retained for future interviewing after the sample was dropped in 1985; 183 of these respondents were interviewed in 1990.

1.4 Multiple Respondent Households

The NLSY79 sample design selected as respondents individuals aged 14 to 21 as of December 31, 1978, who were connected to a surveyed household during 1978. Respondents interviewed in 1979 originated from 8,770 unique households; 2,862 households included more than one NLSY79 respondent.

The most common relationships between respondents living in multiple respondent households at the time the survey began were those of sibling or spouse (see Table 1.4.1). During the 1979 survey, 5,863

respondents were members of a household containing multiple interviewed siblings. More than 330 respondents were members of a household in which their spouse was also interviewed.

**Table 1.4.1 Number of NLSY79 Civilian Respondents
by Type of Household: 1979**

Type of Household	Number of Respondents	Number of Households
Single Respondent	5908	5908
Multiple Siblings		
Two Siblings	3386	1693
Three Siblings	1725	575
Four Siblings	604	151
Five Siblings	130	26
Total Multiple Siblings	5863	2448
Spouse	334	167
Other	581	247
Totals	12686	8770

Note: Siblings may be biological, step, or adopted. Some households may include both siblings and spouses, as well as respondents with other relationships not presented in this table.

1.5 NLSY79 Children

The child sample began in 1986; the expanded mother-child data collection has occurred biennially since then. The child sample includes children born to female NLSY79 respondents. The number of children born to interviewed mothers has increased from 5,255 in 1986 to 8,323 in 2002. Interviews were completed during 2002 with 7,467 children, or 90 percent of children born to interviewed NLSY79 mothers.

Starting with the 1994 survey, the children are treated as two separate groups. The first includes children who were under age 15 (as of December 31 of the survey year). These children are directly assessed (i.e. completed one or more of the assessment instruments), and information about the child is obtained from the child's mother. The second group comprises NLSY79 children who are at least 15 years of age by the end of the survey's calendar year. These "young adults," most of whom were assessed during earlier child surveys, are administered a separate NLSY79-style questionnaire, that gathers information on a wide range of topics.

Sample sizes for NLSY79 mothers and children across survey years are presented in Table 1.5.1. Important issues related to changes over time in the NLSY79 child and mother sample sizes are discussed in the *NLSY79 Child and Young Adult Data Users Guide*, *NLSY79 Children 1992: Description and Evaluation*, and the *NLSY79 Child Handbook: A Guide to the 1986–1990 NLSY79 Child Data*.

Table 1.5.1 NLSY79 Mother & Child Samples: 1979–2002 Surveys

	1979	1986	1988	1990	1992	1994	1996	1998	2000	2002
NLSY79 Female Respondents										
Interviewed	6283	5418	5312	4510	4535	4480	4361	4299	4113	3955
NLSY79 Female Respondents with Children										
# Interviewed	—	2922	3346	3088	3325	3464	3489	3533	3425	3315
# Interviewed who also had one or more children interviewed	—	2774	3196	2772	2964	3212	3228	3221	2934	3122
Children of NLSY79 Female Respondents										
Total Born										
Born to Interviewed Mothers	—	5255	6543	6427	7255	7862	8125	8395	8323	8100
Under Age 15 as of December 31 ⁴						6622	6010	5343	4438	3502
Age 15 and over as of December 31						1240	2113	3052	3885	4598
Total Interviewed ¹	—	4971	6266	5803	6509	7089	7103	7067 ²	6417	7467 ³
Under Age 15 as of December 31						6109	5431	4924	3392	3229
Age 15 and over as of December 31						980	1672	2143	3025	4238

Note: Sample sizes for all child survey years exclude the 441 female members of the military subsample dropped from interviewing in 1985 and the children born to these women. In addition, sample sizes for 1990 and later surveys exclude female members of the civilian economically disadvantaged, non-black/non-Hispanic subsample, whose children were not eligible for assessment. The exclusion of this sample after 1998 accounts for much of the drop in sample size between 1988 and 1990.

¹ An interview was considered complete if an interviewer was able to directly assess a child, or to obtain information from the mother on the child's background and health.

² This total includes 37 children who were assessed or interviewed whose mothers were **not** interviewed.

³ This total includes 14 children (age 0-14) who were assessed or interviewed whose mothers were **not** interviewed and 257 young adult children whose mothers were not interviewed.

⁴ This number includes deceased and non resident children whose mothers were not interviewed.

The child sample interviewed during 2002 includes substantial numbers of Hispanic and black children and children from birth through their late twenties. While the majority of children in the sample had been under 10 years of age in all survey years through 1994, the 1996 survey saw the number of children in the initial sample who are 10 years or older become the majority for the first time in the survey's history. By 2002, the number of young adult children exceeds the number of children under age 15. More than 1700 of these respondents were age 21 and older as of the end of the most recent survey. Table 1.5.2 presents, by select age ranges and race/ethnicity, the number of children across survey years for whom interviews were completed.

Table 1.5.2 NLSY79 Child Sample Sizes by Age & Race/Ethnicity: 1986–2002

	1986	1988	1990	1992	1994	1996	1998	2000	2002
Total Interviewed	4971	6266	5803	6509	7089	7103	7067	6417	7467
By Age¹									
Birth to 9 Years	4676	5380	4508	4430	4154	3480	2978	1912	1619
10 to 14 Years	294	851	1158	1700	2084	1951	1996	1480	1596
15 Years & Older	1	35	137	379	851	1672	2143	3025	4238
By Race/Ethnicity²									
Hispanic	937	1158	1304	1483	1546	1520	1550	1193	1625
Black	1604	1895	1994	2133	2350	2330	2229	1914	2412
Non-black/non-Hispanic	2430	3213	2505	2893	3193	3253	3288	3310	3430

¹ Starting in 1994, children who are aged 15 by December 31st of the interview year are interviewed as young adults. They are counted in this table as 15 years and older though the chronological age at the time of the actual interview might be 14 years for some respondents. Age for children under the age of 15 refers to their age at their mother's interview date.

² Decision rules used in assigning race/ethnicity are described in the *NLSY79 User's Guide*.

The overall sample of children interviewed in 2002 represents a cross-section of children born to women who were 35–43 years of age in 2002. The large number of children born to female respondents range in age from birth to 29 years. It is estimated that the number of children in the sample represents a majority of the children to be born to this contemporary cohort of American women. When weighted, the number of children born to these women can be considered fully representative of children born to a nationally representative sample of women who were between the ages of 14 and 21 on December 31, 1978.

1.6 Special Data Collections

The NLSY79 survey combines respondent interviews with round-specific household enumerations and a series of separately fielded administrative data collections. Annual or biennial surveys collect information directly from NLSY79 sample members about their work, family, and other life course experiences. Round-specific household enumerations describe the composition of the household unit and the characteristics of each household member on the date of the interview. A set of administrative data collections has transcribed data related to NLSY79 sample members found in internal school records. These separately fielded collections include:

- (1) *School Characteristic Information* provided by secondary school administrators during 1980 for the 1979 NLSY79 sample group. School characteristic data were collected during 1995 and 1996 from the primary and secondary schools attended by Children of the NLSY79.
- (2) *School Transcripts*, including coursework and attendance records. Transcripts were collected and coded during the early 1980s, with some amount of data available for 8,951 NLSY79 respondents. School record information was collected during 1995 and 1996 for Children of the NLSY79 who were five years of age and older.
- (3) *Aptitude and Achievement Scores* from standardized tests administered during a young person's schooling. Scores were transcribed from school records between 1980 and 1983 for NLSY79 respondents and during 1995 and 1996 for the Children of the NLSY79.

1.7 NLSY79 Data Files

Almost all NLSY79 data are currently available to the public online or, at a nominal charge, you may purchase the cross-cohort CD which contains the cumulative longitudinal record of each respondent from 1979 to the most recent interview date. In addition to the main NLSY79 data set, special data files also exist. Brief descriptions of the current releases of the NLSY79 appear in the following table. Subsequent data releases will be announced in *NLS News*, the quarterly NLS newsletter, which is available online at <http://www.bls.gov/nls/>.

The cross-cohort CD also includes the most recent data from the NLSY97 and the Original Cohorts. For more information on the contents of the cross-cohort CD consult the *NLS Handbook*.

Table 1.7.1 NLSY79: Current Data Releases

NLSY79 Main Files	1979–2002	The longitudinal record of each respondent including information from the separately fielded 1980 high school survey, 1980 ASVAB administration, and 1980–83 transcript data collections. Data on the NLSY79 main file are supplemented by the four special data files described below.
NLSY79 Child/YA File	1986–2002	Information from a battery of child cognitive, socio-emotional, and physiological assessments administered to NLSY79 mothers and their children during the 1986, 1988, 1990, 1992, 1994, 1996, 1998, 2000 and 2002 NLSY79 surveys, the <i>Child and Mother Supplements</i> , and the <i>Young Adult Survey</i> have been combined with other data collected during the main youth surveys. This data set permits examination of the links between maternal–family behaviors or attitudes and the subsequent development of more than 10,000 children.
NLSY79 Work History	1979–2002	Constructed from work experience data collected during the main NLSY79 surveys, this file provides a week-by-week longitudinal work record of each respondent from January 1, 1978, through the current survey date. It is arranged in three primary arrays: (a) the <i>STATUS</i> array of the respondent's labor force/military status; (b) an <i>HOURLY</i> array of the usual hours worked per week at all jobs; and (c) a <i>DUALJOB</i> array containing additional job numbers for respondents who worked at more than one job simultaneously in any week. In addition, there are 21 years of job-specific data for up to five jobs.
NLSY79 Geocode Files	1979–2002	Information on state, county, and SMSA/MSA/CMSA/PMSA of respondents' current residence, location of most recent college attended, and select environmental variables from the <i>County and City Data Books</i> for county or SMSA of current residence is available to persons whose research work is related to the National Longitudinal Surveys and who satisfactorily complete the Bureau of Labor Statistics' geocode accessing agreement procedure.

1.8 Organization of the Guide

The remainder of the guide is organized as follows:

Chapter 2 contains the technical information on the NLSY79. It includes information on sample sizes and retention rates, sampling design and fielding procedures, sample representativeness and attrition, and weighting and design effects.

Chapter 3 provides information on how NLS variables are collected, created, and arranged in the data files. This chapter details how users can access variables and explains the accompanying documentation.

Chapter 4 presents summary discussions of variable sets arranged alphabetically by topic. Persons interested in reviewing, for example, variables that contain information on a respondent's labor market status or geographic residence should turn to the respective topical section with that name. Each topical section includes variable summaries, references to relevant survey instruments or documentation items, cautionary notes to users about inconsistencies in the data, and helpful hints on how to use the data.

Chapter 5 presents information and explanations of the different types of item nonresponse occurring in the NLSY79. It also outlines the significance of incidences of item nonresponse.

Appendices are the sixth section of the *NLSY79 User's Guide*. These appendices contain a quick reference guide with useful NLS information, a glossary of NLSY79 terms, summaries of the content of the various topical NLSY79 areas of interest, and information on standard errors and design effects.

